Real-time Image detection of cocoa pods in natural environment using deep learning algorithms

Juan Pablo Rojas Bustos[1], Sabine-Karen Lammoglia[2,3], Philippe Borianne[4,5], Frédéric Théveny[4,5], Llorenç Cabrera-Bosquet[6]

Rationale

Estimation of crop yield along its different growth stages is essential when making decisions about disease management, harvest, storage, transport, and marketing. In the case of cacao (*Theobroma cacao*) yield estimation is typically based on manual counts, which are time-consuming, expensive, and often associated to huge estimation errors. Designing a low-cost machine vision system with strong operability for the real-time identification of cocoa pods in different growth stages under natural environments is of great significance for research and cocoa industry. The main goal of this study was to propose a simple and reliable method for cocoa pods segmentation and counting from RGB images acquired with mobile phones, using deep learning techniques.

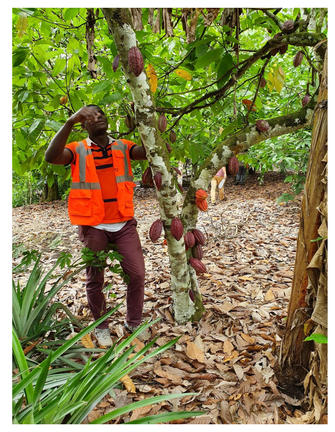
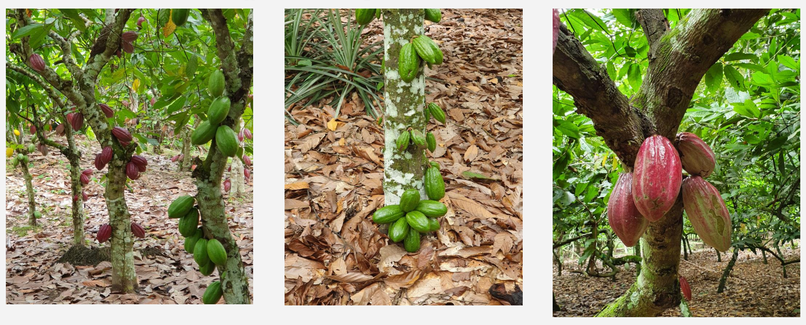


Figure 1. Manual pod counting

Methods

The method presented here involves four steps (Fig. 2), namely (i) image acquisition, (ii) image annotation, (iii) training of models [preprocess and apply training techniques to 80% of the data] with different architectures and (iv) model validation.

Step 1: Image Acquisition (794 RGB images)



Step 2: Image Annotation (*Zooinverse* collaborative platform)

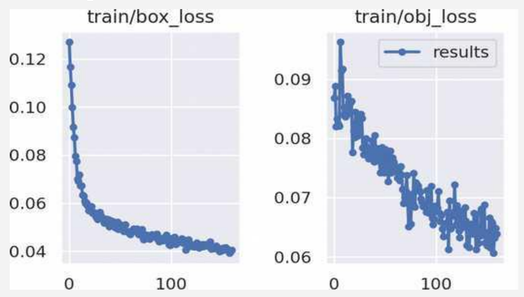
Steps 3: Model training

Yolov5

Faster-RCNN

Training set of **10 955** cocoa annotations (80% of total dataset)

Steps 4: Model validation



Loss curve of Yolov5S, showing model performance

Figure 2. Different steps of the pipeline

Results

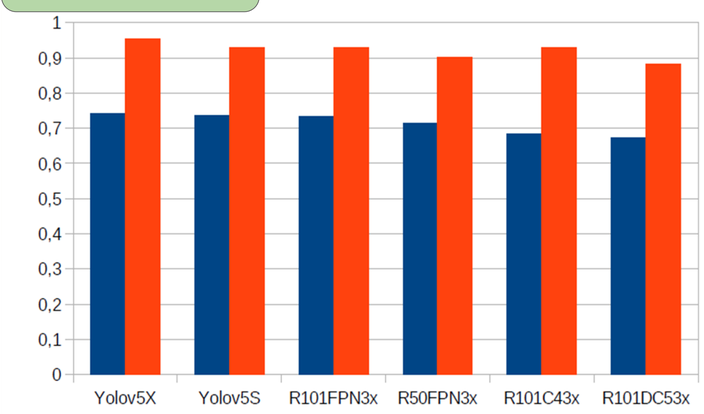


Figure 3. F1 scores of the different tested architectures. Red and blue bars correspond to foreground trees and full images containing trees in the background, respectively.

F1 scores of the different tested architectures for different validation sets

	Green pods detection		Red pods detection	
	Full images	**Foreground trees**	**Full images**	**Foreground trees**
Yolov5X	**0.943**	**0.958**	**0.93**	**0.952**
Yolov5S	0.852	0.92	0.784	0.909
R101FPN$_{3x}$	0.893	0.955	0.608	0.909
R50FPN$_{3x}$	**0.941**	**0.996**	0.625	**0.952**
R101C4$_{3x}$	**0.958**	**1**	0.682	**0.952**
R101DC5$_{3x}$	0.869	0.93	0.684	0.909

- All six networks have difficulty in reliably detecting heterogeneous (shape, color) background pods
- However, with an **F1-score above 95%**, the **Yolov5x, Faster R50FPN3x and R101C43x** give sufficiently close predictions in the foreground trees to be considered equivalent
- Since **Yolov5S** has a simplified architecture that can easily be embedded in smartphones and tablets, its performances can be considered acceptable

Figure 4. Examples of cocoa pod detection (Yolov5x)

Conclusions and perspectives

- The tested models demonstrated a better ability than human operator to detect pods, particularly in the foreground trees
- Future work will focus on (i) establishing a new protocol for image acquisition that will allow only one cocoa tree per image, (ii) improving the detection accuracy, (iii) establishing a robust yield model at the tree level, while also considering the agropedoclimatic factors responsible for the variability of cocoa yields
- A reliable real-time model to detect cocoa pod may have large impacts on the efficiency of cacao management and production profitability. It can help inform cacao farmers early about the expected yield in order to assist in managing the correct fertiliser and pesticide application rate and timing

[1]CIRAD, UMR AGAP, F-34398 Montpellier, France; [2]UFR Biosciences, Université Félix Houphouët-Boigny, Abidjan, Côte d'Ivoire; [3]CIRAD, UMR ABSYS, Montpellier, France; [4]CIRAD, UMR AMAP, F-34398 Montpellier, France; [5]AMAP, Univ Montpellier, CIRAD, CNRS, INRAE, IRD, Montpellier, France; [6]LEPSE, Univ Montpellier, INRAE, Montpellier SupAgro, Montpellier, France